```
Qy       1845 TTTGGCCATTTGTATGCATGATGTTTGGTTTTTAAACATGGTATAATGAATTGTGTACTT 1904
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TTTGGCCATTTGTATGCATGATGTTTGGTTTTTAAACATGGTATAATGAATTGTGTACTT 1920

Qy       1905 CTGTCAGAAGAAAGCAGAGGTACTAATCTCCAATTAAAAAATTTTTTAACATGTAAAAAA 1964
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       1921 CTGTCAGAAGAAAGCAGAGGTACTAATCTCCAATTAAAAAATTTTTTAACATGTAAGAAT 1980

Qy       1965 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1994
                  |     || || || |  | | |||
Db       1981 TTTGTACTTTGAACAACAAGATTACAGAAA 2010
```

RESULT 4
HUM3H3M

LOCUS	HUM3H3M	1650 bp mRNA linear PRI 24-FEB-1995
DEFINITION	Homo sapiens 3-hydroxy-3-methylglutaryl coenzyme A synthase mRNA, complete cds.	
ACCESSION	L25798	
VERSION	L25798.1 GI:410027	
KEYWORDS	3-hydroxy-3-methylglutaryl coenzyme A synthase.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 1650)	
AUTHORS	Rokosz,L.L., Boulton,D.A., Butkiewicz,E.A., Sanyal,G., Cueto,M.A., Lachance,P.A. and Hermes,J.D.	
TITLE	Human cytoplasmic 3-hydroxy-3-methylglutaryl coenzyme A synthase: expression, purification, and characterization of recombinant wild-type and Cys129 mutant enzymes	
JOURNAL	Arch. Biochem. Biophys. 312 (1), 1-13 (1994)	
MEDLINE	94304197	
PUBMED	7913309	
COMMENT	Original source text: Homo sapiens fetal adrenal cDNA to mRNA.	
FEATURES	Location/Qualifiers	
source	1. .1650	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/tissue_type="fetal adrenal"	
CDS	22. .1584	
	/codon_start=1	
	/product="3-hydroxy-3-methylglutaryl coenzyme A synthase"	
	/protein_id="AAA62411.1"	
	/db_xref="GI:410028"	
	/translation="MPGSLPLNAEACWPKDVGIVALEIYFPSQYVDQAELEKYDGVDA GKYTIGLGQAKMGFCTDREDINSLCMTVVQNLMERNNLSYDCIGRLEVGTETIIDKSK SVKTNLMQLFEESGNTDIEGIDTTNACYGGTAAVFNAVNWIESSSWDGRYALVVAGDI AVYATGNARPTGGVGAVALLIGPNAPLIFERGLRGTHMQHAYDFYKPDMLSEYPIVDG KLSIQCYLSALDRCYSVYCKKIHAQWQKEGNDKDFTLNDFGFMIFHSPYCKLVQKSLA RMLLNDFLNDQNRDKNSIYSGLEAFGDVKLEDTYFDRDVEKAFMKASSELFSQKTKAS LLVSNQNGNMYTSSVYGSLASVLAQYSPQQLAGKRIGVFSYGSGLAATLYSLKVTQDA TPGSALDKITASLCDLKSRLDSRTGVAPDVFAENMKLREDTHHLVNYIPQGSIDSLFE GTWYLVRVDEKHRRTYARRPTPNDDTLDEGVGLVHSNIATEHIPSPAKKVPRLPATAA EPEAAVISNGEH"	

ORIGIN

Query Match 68.4%; Score 1370; DB 9; Length 1650;
Best Local Similarity 92.0%; Pred. No. 1e-287;
Matches 1509; Conservative 0; Mismatches 5; Indels 126; Gaps 1;

```
Qy        134 GCTCTTTCACCATGCCTGGATCACTTCCTTTGAATGCAGAAGCTTGCTGGCCAAAAGATG 193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 GCTCTTTCACCATGCCTGGATCACTTCCTTTGAATGCAGAAGCTTGCTGGCCAAAAGATG 70

Qy        194 TGGGAATTGTTGCCCTTGAGATCTATTTTCCTTCTCAATATGTTGATCAAGCAGAGTTGG 253
              | || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db.        71 TTGGGATTGTTGCCCTTGAGATCTATTTTCCTTCTCAATATGTTGATCAAGCAGAGTTGG 130

Qy        254 AAAAATATGATGGTGTAGATGCTGGAAAGTATACCATTGGCTTGGGCCAGGCCAAGATGG 313
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
          131 AAAAATATGATGGTGTAGATGCTGGGAAGTATACCATTGGCTTGGGCCAGGCCAAGATGG 190
```

```
Qy        314 GCTTCTGCACAGATAGAGAAGATATTAACTCTCTTTGCATGACTGTGGTTCAGAATCTTA 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        191 GCTTCTGCACAGATAGAGAAGATATTAACTCTCTTTGCATGACTGTGGTTCAGAATCTTA 250

Qy        374 TGGAGAGAAATAACCTTTCCTATGATTGCATTGGGCGGCTGGAAGTTGGAACAGAGACAA 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 TGGAGAGAAATAACCTTTCCTATGATTGCATTGGGCGGCTGGAAGTTGGAACAGAGACAA 310

Qy        434 TCATCGACAAATCAAAGTCTGTGAAGACTAATTTGATGCAGCTGTTTGAAGAGTCTGGGA 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 TCATCGACAAATCAAAGTCTGTGAAGACTAATTTGATGCAGCTGTTTGAAGAGTCTGGGA 370

Qy        494 ATACAGATATAGAAGGAATCGACACAACTAATGCATGCTATGGAGGCACAGCTGCTGTCT 553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        371 ATACAGATATAGAAGGAATCGACACAACTAATGCATGCTATGGAGGCACAGCTGCTGTCT 430

Qy        554 TCAATGCTGTTAACTGGATTGAGTCCAGCTCTTGGGAT---------------------- 591
              ||||||||||||||||||||||||||||||||||||
Db        431 TCAATGCTGTTAACTGGATTGAGTCCAGCTCTTGGGATGGACGGTATGCCCTGGTAGTTG 490

Qy        592 ------------------------------------------------------------ 591

Db        491 CAGGAGATATTGCTGTATATGCCACAGGAAATGCTAGACCTACAGGTGGAGTTGGAGCAG 550

Qy        592 ----------------------------------------GGGCTTCGTGGGACAC 607
                                                     ||||||||||||||||
Db        551 TAGCTCTGCTAATTGGGCCAAATGCTCCTTTAATTTTTGAACGAGGGCTTCGTGGGACAC 610

Qy        608 ATATGCAACATGCCTATGATTTTTACAAGCCTGATATGCTATCTGAATATCCTATAGTAG 667
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 ATATGCAACATGCCTATGATTTTTACAAGCCTGATATGCTATCTGAATATCCTATAGTAG 670

Qy        668 ATGGAAAACTCTCCATACAGTGCTACCTCAGTGCATTAGACCGCTGCTACTCTGTCTACT 727
              |||| ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        671 ATGGGAAACTCTCCATACAGTGCTACCTCAGTGCATTAGACCGCTGCTATTCTGTCTACT 730

Qy        728 GCAAAAAGATCCATGCCCAGTGGCAGAAAGAGGGAAATGATAAAGATTTTACCTTGAATG 787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        731 GCAAAAAGATCCATGCCCAGTGGCAGAAAGAGGGAAATGATAAAGATTTTACCTTGAATG 790

Qy        788 ATTTTGGCTTCATGATCTTTCACTCACCATATTGTAAACTGGTTCAGAAATCTCTAGCTC 847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        791 ATTTTGGCTTCATGATCTTTCACTCACCATATTGTAAACTGGTTCAGAAATCTCTAGCTC 850

Qy        848 GGATGTTGCTGAATGACTTCCTTAATGACCAGAATAGAGATAAAAATAGTATCTATAGTG 907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        851 GGATGTTGCTGAATGACTTCCTTAATGACCAGAATAGAGATAAAAATAGTATCTATAGTG 910

Qy        908 GCCTGGAAGCCTTTGGGGATGTTAAATTAGAAGACACCTACTTTGATAGAGATGTGGAGA 967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        911 GCCTGGAAGCCTTTGGGGATGTTAAATTAGAAGACACCTACTTTGATAGAGATGTGGAGA 970

Qy        968 AGGCATTTATGAAGGCTAGCTCTGAACTCTTCAGTCAGAAAACAAAGGCATCTTTACTTG 1027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        971 AGGCATTTATGAAGGCTAGCTCTGAACTCTTCAGTCAGAAAACAAAGGCATCTTTACTTG 1030

Qy       1028 TATCAAATCAAAATGGAAATATGTACACATCTTCAGTATATGGTTCCCTTGCATCTGTTC 1087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1031 TATCAAATCAAAATGGAAATATGTACACATCTTCAGTATATGGTTCCCTTGCATCTGTTC 1090

Qy       1088 TAGCACAGTACTCACCTCAGCAATTAGCAGGGAAGAGAATTGGAGTGTTTTCTTATGGTT 1147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1091 TAGCACAGTACTCACCTCAGCAATTAGCAGGGAAGAGAATTGGAGTGTTTTCTTATGGTT 1150

Qy       1148 CTGGTTTGGCTGCCACTCTGTACTCTCTTAAAGTCACACAAGATGCTACACCGGGGTCTG 1207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1151 CTGGTTTGGCTGCCACTCTGTACTCTCTTAAAGTCACACAAGATGCTACACCGGGGTCTG 1210

Qy       1208 CTCTTGATAAAATAACAGCAAGTTTATGTGATCTTAAATCAAGGCTTGATTCAAGAACTG 1267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1211 CTCTTGATAAAATAACAGCAAGTTTATGTGATCTTAAATCAAGGCTTGATTCAAGAACTG 1270
```

```
Qy       1268 GTGTGGCACCAGATGTCTTCGCTGAAAACATGAAGCTCAGAGAGGACACCCATCATTTGG 1327
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1271 GTGTGGCACCAGATGTCTTCGCTGAAAACATGAAGCTCAGAGAGGACACCCATCATTTGG 1330

Qy       1328 TCAACTATATTCCCCAGGGTTCAATAGATTCACTCTTTGAAGGAACGTGGTACTTAGTTA 1387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1331 TCAACTATATTCCCCAGGGTTCAATAGATTCACTCTTTGAAGGAACGTGGTACTTAGTTA 1390

Qy       1388 GGGTGGATGAAAAGCACAGAAGAACTTACGCTCGGCGTCCCACTCCAAATGATGACACTT 1447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1391 GGGTGGATGAAAAGCACAGAAGAACTTACGCTCGGCGTCCCACTCCAAATGATGACACTT 1450

Qy       1448 TGGATGAAGGAGTAGGACTTGTGCATTCAAACATAGCAACTGAGCATATTCCAAGCCCTG 1507
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1451 TGGATGAAGGAGTAGGACTTGTGCATTCAAACATAGCAACTGAGCATATTCCAAGCCCTG 1510

Qy       1508 CCAAGAAAGTACCAAGACTCCCTGCCACAGCAGCAGAACCTGAAGCAGCTGTCATTAGTA 1567
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1511 CCAAGAAAGTACCAAGACTCCCTGCCACAGCAGCAGAACCTGAAGCAGCTGTCATTAGTA 1570

Qy       1568 ATGGGGAACATTAAGATACTCTGTGAGGTGCAAGACTTCAGGGTGGGGTGGGCATGGGGT 1627
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1571 ATGGGGAACATTAAGATACTCTGTGAGGTGCAAGACTTCAGGGTGGGGTGGGCATGGGGT 1630

Qy       1628 GGGGGTATGGGAACAGTTGG 1647
             ||||||||||||||||||||
Db       1631 GGGGGTATGGGAACAGTTGG 1650
```

```
RESULT 5
BT007302
LOCUS       BT007302                1563 bp    mRNA    linear   PRI 13-MAY-2003
DEFINITION  Homo sapiens 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1
            (soluble) mRNA, complete cds.
ACCESSION   BT007302
VERSION     BT007302.1  GI:30583442
KEYWORDS    FLI_CDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1563)
  AUTHORS   Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
            Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
            Phelan,M. and Farmer,A.
  TITLE     Cloning of human full-length CDSs in BD Creator(TM) System Donor
            vector
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1563)
  AUTHORS   Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
            Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
            Phelan,M. and Farmer,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
            Circle, Palo Alto, CA 94303, USA
COMMENT     This CDS clone is a part of a collection of human full length
            expression clones generated by BD Biosciences Clontech and the
            Harvard Institute of Proteomics. Each CDS has been cloned in two
            forms: with and without stop-codon (to allow fusion with C-terminal
            tag). The CDS has been  directionally cloned using BD In-Fusion(TM)
            cloning system between the SalI  and HindIII sites of the pDNR-DUAL
            vector. Additional sequences in the clone:  'ACC' after SalI site
            and before 'ATG' to provide Kozak consensus sequence; 'GG' after
            last codon and before HindIII site to maintain reading frame.
            Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES             Location/Qualifiers
     source          1. .1563
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="GH00878X1.0"
                     /clone_lib="BD Creator(TM) CDS Library derived from MGC
                     collection"
                     /lab_host="DH5alpha T1 resistant"
                     /note="Vector: pDNR-Dual"
     CDS             1. .1563
                     /codon_start=1
                     /product="3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1
                     (soluble)"
                     /protein_id="AAP35966.1"
                     /db_xref="GI:30583443"
                     /translation="MPGSLPLNAEACWPKDVGIVALEIYFPSQYVDQAELEKYDGVDA
                     GKYTIGLGQAKMGFCTDREDINSLCMTVVQNLMERNNLSYDCIGRLEVGTETIIDKSK
                     SVKTNLMQLFEESGNTDIEGIDTTNACYGGTAAVFNAVNWIESSSWDGRYALVVAGDI
                     AVYATGNARPTGGVGAVALLIGPNAPLIFERGLRGTHMQHAYDFYKPDMLSEYPIVDG
                     KLSIQCYLSALDRCYSVYCKKIHAQWQKEGNDKDFTLNDFGFMIFHSPYCKLVQKSLA
                     RMLLNDFLNDQNRDKNSIYSGLEAFGDVKLEDTYFDRDVEKAFMKASSELFSQKTKAS
                     LLVSNQNGNMYTSSVYGSLASVLAQYSPQQLAGKRIGVFSYGSGLAATLYSLKVTQDA
                     TPGSALDKITASLCDLKSRLDSRTGVAPDVFAENMKLREDTHHLVNYIPQGSIDSLFE
                     GTWYLVRVDEKHRRTYARRPTPNDDTLDEGVGLVHSNIATEHIPSPAKKVPRLPATAA
                     EPEAAVISNGEH"
ORIGIN
```

```
  Query Match           64.9%;  Score 1300;  DB 9;  Length 1563;
  Best Local Similarity   91.9%;  Pred. No. 1.7e-272;
  Matches 1436;  Conservative    0;  Mismatches    0;  Indels  126;  Gaps    1;

Qy        145 ATGCCTGGATCACTTCCTTTGAATGCAGAAGCTTGCTGGCCAAAAGATGTGGGAATTGTT 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCTGGATCACTTCCTTTGAATGCAGAAGCTTGCTGGCCAAAAGATGTGGGAATTGTT 60

Qy        205 GCCCTTGAGATCTATTTTCCTTCTCAATATGTTGATCAAGCAGAGTTGGAAAAATATGAT 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCTTGAGATCTATTTTCCTTCTCAATATGTTGATCAAGCAGAGTTGGAAAAATATGAT 120

Qy        265 GGTGTAGATGCTGGAAAGTATACCATTGGCTTGGGCCAGGCCAAGATGGGCTTCTGCACA 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTGTAGATGCTGGAAAGTATACCATTGGCTTGGGCCAGGCCAAGATGGGCTTCTGCACA 180

Qy        325 GATAGAGAAGATATTAACTCTCTTTGCATGACTGTGGTTCAGAATCTTATGGAGAGAAAT 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATAGAGAAGATATTAACTCTCTTTGCATGACTGTGGTTCAGAATCTTATGGAGAGAAAT 240

Qy        385 AACCTTTCCTATGATTGCATTGGGCGGCTGGAAGTTGGAACAGAGACAATCATCGACAAA 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACCTTTCCTATGATTGCATTGGGCGGCTGGAAGTTGGAACAGAGACAATCATCGACAAA 300

Qy        445 TCAAAGTCTGTGAAGACTAATTTGATGCAGCTGTTTGAAGAGTCTGGGAATACAGATATA 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAAAGTCTGTGAAGACTAATTTGATGCAGCTGTTTGAAGAGTCTGGGAATACAGATATA 360

Qy        505 GAAGGAATCGACACAACTAATGCATGCTATGGAGGCACAGCTGCTGTCTTCAATGCTGTT 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAAGGAATCGACACAACTAATGCATGCTATGGAGGCACAGCTGCTGTCTTCAATGCTGTT 420

Qy        565 AACTGGATTGAGTCCAGCTCTTGGGAT-------------------------------- 591
              |||||||||||||||||||||||||||
Db        421 AACTGGATTGAGTCCAGCTCTTGGGATGGACGGTATGCCCTGGTAGTTGCAGGAGATATT 480

Qy        592 ---------------------------------------------------------- 591

Db        481 GCTGTATATGCCACAGGAAATGCTAGACCTACAGGTGGAGTTGGAGCAGTAGCTCTGCTA 540

Qy        592 --------------------------------GGGCTTCGTGGGACACATATGCAACAT 618
                                               ||||||||||||||||||||||||||||
Db        541 ATTGGGCCAAATGCTCCTTTAATTTTTGAACGAGGGCTTCGTGGGACACATATGCAACAT 600

Qy        619 GCCTATGATTTTTACAAGCCTGATATGCTATCTGAATATCCTATAGTAGATGGAAAACTC 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCCTATGATTTTTACAAGCCTGATATGCTATCTGAATATCCTATAGTAGATGGAAAACTC 660

Qy        679 TCCATACAGTGCTACCTCAGTGCATTAGACCGCTGCTACTCTGTCTACTGCAAAAAGATC 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCCATACAGTGCTACCTCAGTGCATTAGACCGCTGCTACTCTGTCTACTGCAAAAAGATC 720

Qy        739 CATGCCCAGTGGCAGAAAGAGGGAAATGATAAAGATTTTACCTTGAATGATTTTGGCTTC 798
```

OM protein - protein search, using sw model

Run on: June 24, 2004, 12:56:26 ; Search time 44 Seconds
(without alignments)
565.671 Million cell updates/sec

Title: US-10-622-516-2
Perfect score: 2511
Sequence: 1 MPGSLPLNAEACWPKDVGIV.....PRLPATAAEPEAAVISNGEH 478

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2480	98.8	520	1	HMCS_HUMAN	Q01581 homo sapien
2	2362	94.1	520	1	HMCS_RAT	P17425 rattus norv
3	2352	93.7	520	1	HMCS_CRIGR	P13704 cricetulus
4	2097	83.5	522	1	HMCS_CHICK	P23228 gallus gall
5	1513	60.3	508	1	HMCM_HUMAN	P54868 homo sapien
6	1494	59.5	508	1	HMCM_RAT	P22791 rattus norv
7	1493	59.5	480	1	HMCM_MOUSE	P54869 mus musculu
8	1466	58.4	508	1	HMCM_PIG	O02734 sus scrofa
9	1352.5	53.9	453	1	HMC1_BLAGE	P54961 blattella g
10	1272.5	50.7	455	1	HMC2_BLAGE	P54870 blattella g
11	980	39.0	461	1	HMCS_ARATH	P54873 arabidopsis
12	950.5	37.9	491	1	HMCS_YEAST	P54839 saccharomyc
13	900	35.8	447	1	HMCS_SCHPO	P54874 schizosacch
14	705	28.1	462	1	HMCS_CAEEL	P54871 caenorhabdi
15	249	9.9	163	1	HMCS_DICDI	P54872 dictyosteli
16	152	6.1	350	1	YD79_METKA	Q8tvl0 methanopyru
17	147	5.9	345	1	PKSG_BACSU	P40830 bacillus su
18	145.5	5.8	346	1	Y792_METTH	O26883 methanobact
19	144	5.7	350	1	Y677_PYRHO	O58410 pyrococcus
20	138.5	5.5	350	1	Y972_PYRFU	Q51798 pyrococcus
21	135	5.4	349	1	Y871_METMA	Q8pyj0 methanosarc
22	132.5	5.3	343	1	YO15_ARCFU	O30256 archaeoglob
23	130	5.2	350	1	YL85_PYRAE	Q8zvp4 pyrobaculum
24	127	5.1	349	1	Y4E1_METAC	Q8tiv0 methanosarc
25	126.5	5.0	350	1	YD69_PYRAB	Q9uyy8 pyrococcus
26	123.5	4.9	716	1	BGAL_THETU	P26257 thermoanaer
27	123	4.9	351	1	Y132_THEVO	Q97cg9 thermoplasm
28	122	4.9	345	1	YF46_METJA	Q58941 methanococc
29	122	4.9	351	1	YE55_THEAC	Q9hi87 thermoplasm
30	121	4.8	580	1	NO56_MOUSE	Q9d6z1 mus musculu
31	118.5	4.7	596	1	NO56_HUMAN	O00567 homo sapien
32	116	4.6	348	1	YD49_SULTO	Q971k8 sulfolobus
33	107.5	4.3	776	1	TOP1_RICPR	Q9zdk2 rickettsia
34	106.5	4.2	1517	1	GLTB_ECOLI	P09831 escherichia
35	106	4.2	367	1	YQL2_CAEEL	Q09291 caenorhabdi
36	105	4.2	412	1	AATC_MOUSE	P05201 mus musculu
37	105	4.2	3358	1	PGCV_MOUSE	Q62059 mus musculu
38	103	4.1	489	1	YSV5_CAEEL	Q10011 caenorhabdi
39	103	4.1	1197	1	DPOM_PODAN	Q01529 podospora a
40	102	4.1	360	1	HIS8_LACLA	Q02135 lactococcus
41	102	4.1	1163	1	CQAA_BACTF	Q9x597 bacillus th
42	101	4.0	365	1	YI73_AERPE	Q9yas0 aeropyrum p
43	100.5	4.0	776	1	TOP1_RICCN	Q92ih1 rickettsia
44	100	4.0	1136	1	C4BA_BACTI	P05519 bacillus th
45	99.5	4.0	494	1	ILVC_VIBPA	Q87tn4 vibrio para

ALIGNMENTS

RESULT 1
HMCS_HUMAN
ID HMCS_HUMAN STANDARD; PRT; 520 AA.
AC Q01581;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hydroxymethylglutaryl-CoA synthase, cytoplasmic (EC 2.3.3.10) (HMG-CoA
DE synthase) (3-hydroxy-3-methylglutaryl coenzyme A synthase).
GN HMGCS1 OR HMGCS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=93041939; PubMed=1358203;
RA Russ A.P., Ruzicka V., Maerz W., Appelhans H., Gross W.;
RT "Amplification and direct sequencing of a cDNA encoding human
RT cytosolic 3-hydroxy-3-methylglutaryl-coenzyme A synthase.";
RL Biochim. Biophys. Acta 1132:329-331(1992).
RN [2]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS.
RC TISSUE=Fetal adrenal gland;
RX MEDLINE=94304197; PubMed=7913309;
RA Rokosz L.L., Boulton D.A., Butkiewicz E.A., Sanyal G., Cueto M.A.,
RA Lachance P.A., Hermes J.D.;
RT "Human cytoplasmic 3-hydroxy-3-methylglutaryl coenzyme A synthase:
RT expression, purification, and characterization of recombinant
RT wild-type and Cys129 mutant enzymes.";
RL Arch. Biochem. Biophys. 312:1-13(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```
CC   -!- FUNCTION: This enzyme condenses acetyl-CoA with acetoacetyl-CoA to
CC       form HMG-CoA, which is the substrate for HMG-CoA reductase.
CC   -!- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + acetoacetyl-CoA = (S)-3-
CC       hydroxy-3-methylglutaryl-CoA + CoA.
CC   -!- PATHWAY: PRODUCTION OF MEVALONATE FROM HMG-COA PRIOR TO THE
CC       SYNTHESIS OF STEROLS SUCH AS CHOLESTEROL AND ISOPRENOIDS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the HMG-CoA synthase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X66435; CAA47061.1; -.
DR   EMBL; L25798; AAA62411.1; -.
DR   EMBL; BC000297; AAH00297.1; -.
DR   PIR; S27197; S27197.
DR   PIR; S45497; S45497.
DR   Genew; HGNC:5007; HMGCS1.
DR   MIM; 142940; -.
DR   GO; GO:0005737; C:cytoplasm; TAS.
DR   GO; GO:0005625; C:soluble fraction; TAS.
DR   GO; GO:0004421; F:hydroxymethylglutaryl-CoA synthase activity; TAS.
DR   InterPro; IPR000590; HMG_CoA_synt_AS.
DR   InterPro; IPR008260; HMG_CoA_synth.
DR   Pfam; PF01154; HMG_CoA_synt; 1.
DR   PROSITE; PS01226; HMG_COA_SYNTHASE; 1.
KW   Transferase; Cholesterol biosynthesis; Multigene family.
FT   ACT_SITE    129    129
FT   MUTAGEN     129    129       C->A,S: LOSS OF ACTIVITY.
FT   CONFLICT    248    248       G -> A (IN REF. 1).
FT   CONFLICT    251    251       K -> N (IN REF. 1).
FT   CONFLICT    299    299       E -> K (IN REF. 1).
FT   CONFLICT    364    364       Q -> H (IN REF. 1).
FT   CONFLICT    420    420       P -> Q (IN REF. 1).
FT   CONFLICT    519    520       EH -> VW (IN REF. 1).
SQ   SEQUENCE   520 AA;  57293 MW;  C669212BF86CFF9B CRC64;

  Query Match             98.8%;  Score 2480;  DB 1;  Length 520;
  Best Local Similarity   91.9%;  Pred. No. 2.3e-173;
  Matches  478;  Conservative    0;  Mismatches    0;  Indels   42;  Gaps    1;

Qy          1 MPGSLPLNAEACWPKDVGIVALEIYFPSQYVDQAELEKYDGVDAGKYTIGLGQAKMGFCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGSLPLNAEACWPKDVGIVALEIYFPSQYVDQAELEKYDGVDAGKYTIGLGQAKMGFCT 60

Qy         61 DREDINSLCMTVVQNLMERNNLSYDCIGRLEVGTETIIDKSKSVKTNLMQLFEESGNTDI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DREDINSLCMTVVQNLMERNNLSYDCIGRLEVGTETIIDKSKSVKTNLMQLFEESGNTDI 120

Qy        121 EGIDTTNACYGGTAAVFNAVNWIESSSWD------------------------------ 149
              ||||||||||||||||||||||||||||
Db        121 EGIDTTNACYGGTAAVFNAVNWIESSSWDGRYALVVAGDIAVYATGNARPTGGVGAVALL 180

Qy        150 -----------GLRGTHMQHAYDFYKPDMLSEYPIVDGKLSIQCYLSALDRCYSVYCKKI 198
                          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 IGPNAPLIFERGLRGTHMQHAYDFYKPDMLSEYPIVDGKLSIQCYLSALDRCYSVYCKKI 240

Qy        199 HAQWQKEGNDKDFTLNDFGFMIFHSPYCKLVQKSLARMLLNDFLNDQNRDKNSIYSGLEA 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HAQWQKEGNDKDFTLNDFGFMIFHSPYCKLVQKSLARMLLNDFLNDQNRDKNSIYSGLEA 300

Qy        259 FGDVKLEDTYFDRDVEKAFMKASSELFSQKTKASLLVSNQNGNMYTSSVYGSLASVLAQY 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 FGDVKLEDTYFDRDVEKAFMKASSELFSQKTKASLLVSNQNGNMYTSSVYGSLASVLAQY 360

Qy        319 SPQQLAGKRIGVFSYGSGLAATLYSLKVTQDATPGSALDKITASLCDLKSRLDSRTGVAP 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SPQQLAGKRIGVFSYGSGLAATLYSLKVTQDATPGSALDKITASLCDLKSRLDSRTGVAP 420

Qy        379 DVFAENMKLREDTHHLVNYIPQGSIDSLFEGTWYLVRVDEKHRRTYARRPTPNDDTLDEG 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 DVFAENMKLREDTHHLVNYIPQGSIDSLFEGTWYLVRVDEKHRRTYARRPTPNDDTLDEG 480

Qy        439 VGLVHSNIATEHIPSPAKKVPRLPATAAEPEAAVISNGEH 478
              |||||||||||||||||||||||||||||||||||||||||
Db        481 VGLVHSNIATEHIPSPAKKVPRLPATAAEPEAAVISNGEH 520
```

```
RESULT 2
HMCS_RAT
ID   HMCS_RAT       STANDARD;      PRT;   520 AA.
AC   P17425;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hydroxymethylglutaryl-CoA synthase, cytoplasmic (EC 2.3.3.10) (HMG-CoA
DE   synthase) (3-hydroxy-3-methylglutaryl coenzyme A synthase).
GN   HMGCS1 OR HMGCS.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=90301491; PubMed=1972979;
RA   Ayte J., Gil-Gomez G., Hegardt F.G.;
RT   "Nucleotide sequence of a rat liver cDNA encoding the cytosolic 3-
RT   hydroxy-3-methylglutaryl coenzyme A synthase.";
RL   Nucleic Acids Res. 18:3642-3642(1990).
CC   -!- FUNCTION: This enzyme condenses acetyl-CoA with acetoacetyl-CoA to
CC       form HMG-CoA, which is the substrate for HMG-CoA reductase.
CC   -!- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + acetoacetyl-CoA = (S)-3-
CC       hydroxy-3-methylglutaryl-CoA + CoA.
CC   -!- PATHWAY: PRODUCTION OF MEVALONATE FROM HMG-COA PRIOR TO THE
CC       SYNTHESIS OF STEROLS SUCH AS CHOLESTEROL AND ISOPRENOIDS.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the HMG-CoA synthase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52625; CAA36852.1; -.
DR   PIR; S12736; S12736.
DR   InterPro; IPR000590; HMG_CoA_synt_AS.
DR   InterPro; IPR008260; HMG_CoA_synth.
DR   Pfam; PF01154; HMG_CoA_synt; 1.
DR   PROSITE; PS01226; HMG_COA_SYNTHASE; 1.
KW   Transferase; Cholesterol biosynthesis; Multigene family.
FT   ACT_SITE    129    129       POTENTIAL.
SQ   SEQUENCE   520 AA;  57433 MW;  CB213A27B0C177CB CRC64;

  Query Match             94.1%;  Score 2362;  DB 1;  Length 520;
  Best Local Similarity   87.5%;  Pred. No. 9.2e-165;
  Matches  455;  Conservative   11;  Mismatches   12;  Indels   42;  Gaps    1;

Qy          1 MPGSLPLNAEACWPKDVGIVALEIYFPSQYVDQAELEKYDGVDAGKYTIGLGQAKMGFCT 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db          1 MPGSLPLNAEACWPKDVGIVALEIYFPSQYVDQAELEKYDGVDAGKYTIGLGQARMGFCT 60

Qy         61 DREDINSLCMTVVQNLMERNNLSYDCIGRLEVGTETIIDKSKSVKTNLMQLFEESGNTDI 120
              |||||||||:|||| ||||||:|||||||||||||||||||||:|||||||||||||||
Db         61 DREDINSLCLTVVQKLMERNSLSYDCIGRLEVGTETIIDKSKSVKSNLMQLFEESGNTDI 120

Qy        121 EGIDTTNACYGGTAAVFNAVNWIESSSWD------------------------------ 149
              ||||||||||||||||||||||||||||
```